

Steganalysis of $\pm k$ Steganography based on Noncausal Linear Predictor

K.M. Singh, Y.J. Chanu, T. Tuithung

K. Manglem Singh

Department of Computer Science & Engineering,
NIT Manipur, Imphal, India
Email: manglem@gmail.com

Yambem Jina Chanu*, Themrichon Tuithung

Department of Computer Science & Engineering,
NERIST, Itanagar, India
*Corresponding author: jina.yambem@gmail.com
tth@nerist.ac.in

Abstract: The paper proposes a novel steganalytic technique for $\pm k$ steganography based on noncausal linear predictor using prediction coefficients obtained from the autocorrelation matrix for a block of pixels in the stego-image. The image is divided into equal-size blocks, autocorrelation matrix is found for the block, and the appropriate noncausal linear prediction coefficients is selected to predict all pixels in that block. A pixel is assumed to be embedded with message bit if the absolute difference between the original pixel value and predicted pixel value exceeds the pre-defined threshold. The effectiveness of the proposed technique is verified using different images.

Keywords: LSB embedding, noncausal linear predictor, RS analysis, steganalysis, steganography.

1 Introduction

Our world is becoming smaller and smaller day by day due to the rapid development in computer and network technology, empowered with unlimited expansion of communication and information technology capability globally for exchange of information with large involvement of individuals at home, private and public sectors and government organizations. Both business and society now are highly dependent on the Internet and multimedia technology as an integral part for communication, which on other hand serves two purposes - one for the exchange of vast amount of information and knowledge for the welfare of human beings, and another for the destruction of humanity by using such medium. Law enforcement personnel and criminals conceal their message, work plan, and important information in digital media. This technique is called steganography, which is the art and science of embedding secret messages inside different cover media such as text, audio, image and video without any suspicion. The Internet itself provides hundreds of freeware and shareware, which can be freely used by anyone including criminals and terrorists. To detect and extract the hidden message from stego-data by the cyber security personnel and law enforcement professional is the current interest of steganalysis, which is the science of discovery of existence of hidden information.

The race between steganography and steganalysis is ongoing and never-ending competition, in which the former develops newer and more robust algorithms to embed secret message into cover medium such as text, audio, image and video to form stego-data, which is not different from the original medium perceptually to the ordinary person, and does not arouse any suspicion, while the later tries to stop all steganographic techniques or detect the embedded secret message from the stego-data. Two important media which appeal to hide secret message are the image and video due to the inherent presence of high redundancy in representation of such

data. The most obvious of steganography is to hide data, but its general applications include covert and invisible communication between two or more parties, secret data storing in storage devices, integrity checking of data by using hash value, access control mechanism for digital content distribution by including access key for selective extraction, and media database system by unifying the media data such as music, picture, video with other information like date and time of recording such data as metadata [1].

Steganalysis finds applications in gathering and tracking criminal and cyber terrorist activities and anti-social elements over the Internet, cyber forensics investigation for extraction of hidden message from compromised digital systems, cyber warfare for waging war using the Internet, peaceful purposes for improving steganographic tools to identify their weakness [2]. In the recent decades, cyber security community and professionals take challenging, interesting and exciting researches in developing new and robust steganographic algorithms as well as better steganalytic algorithms to counter the steganography [3–5]. The powerful and popular LSB detection algorithms are Chi-square [6], RS [7], Gradient Energy-Flipping Rate (GEFR) Detection [8] and Histogram difference [9]. These algorithms can estimate the length of the embedded message in LSB embedding. LSB matching (LSBM) steganographic technique adds or subtract by 1 if LSB does not match with message bit [10] [11]. LSBM is a special case of k steganography with $k=1$ [12]. In k steganography, pixel values are either increased or decreased by k . k steganography is an important steganographic technique, and number of steganalytic algorithms for k steganography particularly based on linear prediction is very few. This motivates to develop this novel k steganalytic algorithm.

This paper proposes a new steganalytic technique for k steganography using noncausal linear predictor based on prediction coefficients obtained from the autocorrelation matrix for each block of pixels in the image. Most appropriate elements from that matrix are used for predicting all pixel values in that block. The difference between the predicted pixel value and the original pixel value gives the knowledge about the presence of message bit.

The paper is organized as follows. Section 2 deals with a short introduction on the order of prediction, which is used in the proposed algorithm. Section 3 proposes the new steganalytic algorithm, followed experimental results in Section 4 and conclusions in Section 5.

2 Order of Linear Prediction

The order of linear prediction in images can be based on the neighbourhood set G , which is defined at a coordinate $(m,n) \in G$ with the points lying within a specified radius as its elements [13]. An η^{th} order neighbourhood set at coordinate (m,n) is defined using the Euclidian distance for defining the radius as

$$S_{(m,n)}^\eta = \{(i,j) : 0 < (m-i)^2 + (n-j)^2 \leq D_\eta\} \quad (1)$$

where (i,j) is the coordinates of pixels in the neighbourhood set, $(i,j) \neq (m,n)$, D_η and is an increasing integer function of η . Take for examples

- 1) 1st order neighbourhood set : $\eta = 1; D_\eta = 1;$
 $S_{(m,n)}^1 = \{(m-1, n), (m+1, n), (m, n-1), (m, n+1)\}$
- 2) 2nd order neighbourhood set : $\eta = 2; D_\eta = 2;$
 $S_{(m,n)}^2 = \{(m-1, n), (m+1, n), (m, n-1), (m, n+1), (m-1, n-1),$
 $(m+1, n+1), (m-1, n+1), (m+1, n-1)\}$

Similarly, the 3rd and 4th order neighbourhood sets are defined with $\eta = 3$, $D_\eta = 4$ and $\eta = 4$, $D_\eta = 5$ respectively. The neighbourhood sets up to order 5 for linear prediction is shown in Figure. 1. The pixel to be predicted is labeled '0'. For the first-order prediction, the neighbourhood set consists of pixels marked '1', for the second-order prediction, the neighbourhoods set involves pixels marked as '1' and '2' and so on. Figure. 1 shows the causal and noncausal regions of supports with pixels lying to the left of dark line for the causal prediction.

3 Proposed Steganalytic Algorithm

Linear prediction is used for modelling, estimating and coding of one-dimensional signals, in speech coding and understanding, geophysics and biomedical signal processing applications [14], and of two-dimensional signals in image compression, image segmentation and classification and spectrum estimation [15]. Linear prediction model is used in estimating pixels at any location of the image due to strong correlation among neighbouring pixels in the image. Most of the applications of linear prediction use the causal linear prediction, which predicts the current pixel value based on its surrounding past pixels only. Sometime, same prediction coefficients are used to predict entire pixels for the same block in the image. Asif and Moura wrote that the noncausal linear prediction, which is based on both the past and future pixels in the neighbourhood of the current pixels give better results [13].

(5) $x(m-2, n-2)$	(4) $x(m-2, n-1)$	(3) $x(m-2, n)$	(4) $x(m-2, n+1)$	(5) $x(m-2, n+2)$
(4) $x(m-1, n-2)$	(2) $x(m-1, n-1)$	(1) $x(m-1, n)$	(2) $x(m-1, n+1)$	(4) $x(m-1, n+2)$
(3) $x(m, n-2)$	(1) $x(m, n-1)$	(0) $x(m, n)$	(1) $x(m, n+1)$	(3) $x(m, n+2)$
(4) $x(m+1, n-2)$	(2) $x(m+1, n-1)$	(1) $x(m+1, n)$	(2) $x(m+1, n+1)$	(4) $x(m+1, n+2)$
(5) $x(m+2, n-2)$	(4) $x(m+2, n-1)$	(3) $x(m+2, n)$	(4) $x(m+2, n+1)$	(5) $x(m+2, n+2)$

Figure 1: The neighbourhood sets for the 1st-order to 5th-order prediction.

The predicted pixel $\hat{x}(m, n)$ at the location (m, n) is given by

$$\hat{x}(m, n) = \sum_{(i,j) \in W} a(i, j) x(m-i, n-j) = x_\eta a_\eta \quad (2)$$

where W represents the prediction window that excludes $(0, 0)$ and $a(i, j)$ is the prediction coefficient, x_η is a row vector of pixels used for prediction and a_η is the column vector formed by the corresponding prediction coefficients. The linear prediction error (LPE) is given by

$$e(m, n) = x(m, n) - \hat{x}(m, n) = x(m, n) - x_\eta a_\eta \quad (3)$$

The corresponding mean-square error $E(e^2(m, n))$ is given by

$$E(e^2(m, n)) = E(x(m, n) - x_\eta a_\eta)^2 = E(x^2(m, n)) - 2\acute{a}_\eta E(\acute{x}_\eta x(m, n)) + \acute{a}_\eta E(\acute{a}_\eta x_\eta) a_\eta$$

$$= R_x(0,0) - 2\hat{a}_\eta r_{x_\eta} + \hat{a}_\eta R_{x_\eta} a_\eta \quad (4)$$

where $R_{x_\eta} = E(\hat{x}_\eta x_\eta)$ is the autocorrelation matrix, $r_{x_\eta} = E(x(m,n)\hat{x}_\eta)$ is its cross-correlation vector with the pixel to be predicted and $R_x(0,0) = E(x^2(m,n))$. The optimal prediction coefficient vector is obtained by minimizing $E(e^2(m,n))$ with respect to a_η and found by the condition

$$\nabla(E(e^2(m,n))) = 0 \quad (5)$$

where ∇ is the gradient operator with respect to the vector a_η . This yields in the following matrix form of *normal equations* or *Wiener equations*

$$R_{x_\eta} a_\eta = r_{x_\eta} \quad (6)$$

The solution of the above matrix equation gives the prediction coefficient vector a_η . Using the symmetry property $R_x(i,j) = R_x(-i,-j)$ of the autocorrelation function of wide-sense stationary image random field, it can be shown that [16]

$$a(-i,-j) = a(i,j) \quad (7)$$

The symmetry property simplifies the computation of the prediction coefficients. The predicted pixel value $\hat{x}(m,n)$ should be different from the original pixel value $x(m,n)$ by k (where k is an integer value by which the value of the pixel at (m,n) is changed to hide a binary bit) if the message bit is hidden in the pixel value $x(m,n)$. The absolute error between the original and the predicted pixel values is given by

$$|e(m,n)| = |x(m,n) - \hat{x}(m,n)| \quad (8)$$

The total number of absolute errors, which exceeds a pre-defined threshold Θ is the estimate of length of hidden message bit in the stego-image. The following sub-sections are on different causal and noncausal linear predictors for the first and second order.

3.1 First-order Causal Linear Predictor

The predicted pixel $\hat{x}(m,n)$ of the centre pixel $x(m,n)$ for the first-order causal linear prediction is given by

$$\begin{aligned} \hat{x}(m,n) &= a_1(0,1)x(m,n-1) + a_1(1,0)x(m-1,n) = x_1 a_1 \\ &= [x(m,n-1)x(m-1,n)] \begin{bmatrix} a_1(0,1) \\ a_1(1,0) \end{bmatrix} \end{aligned} \quad (9)$$

The pixels $x(m,n-1)$ and $x(m-1,n)$, marked as "1" are shown to the left side of dark line in Figure.1, and used in the first-order causal linear prediction. The prediction coefficients $a_1(0,1)$ and $a_1(1,0)$ for the first-order causal predictor are related with the autocorrelation functions by the following normal equation in Eq. 6: $R_{x_1} a_1 = r_{x_1}$ where

$$R_{x_1} = \begin{bmatrix} R_x(0,0) & R_x(1,-1) \\ R_x(1,-1) & R_x(0,0) \end{bmatrix} \text{ and } r_{x_1} = \begin{bmatrix} R_x(0,1) \\ R_x(1,0) \end{bmatrix}$$

The number of different elements from the autocorrelation matrix is 4 for the first-order causal predictor.

3.2 First-order Noncausal Linear Predictors

The predicted pixel $\hat{x}(m, n)$ of the centre pixel $x(m, n)$ for the first-order noncausal linear prediction is given by

$$\begin{aligned}\hat{x}(m, n) &= a_1(0, 1)x(m, n-1) + a_1(0, -1)x(m, n+1) + a_1(1, 0)x(m-1, n) + a_1(-1, 0)x(m+1, n) \\ &= a_1(0, 1)x(m, n-1) + x(m, n+1) + a_1(1, 0)x(m-1, n) + x(m+1, n) = x_1 a_1 \\ &= [x(m, n-1) + x(m, n+1) \quad x(m-1, n) + x(m+1, n)]' [a_1(0, 1) \quad a_1(1, 0)]' \quad (10)\end{aligned}$$

where $a_1(0, 1) = a_1(0, -1)$, and $a_1(1, 0) = a_1(-1, 0)$ due to Eq. 7. The pixels $x(m, n-1)$, $x(m, n+1)$, $x(m-1, n)$ and $x(m+1, n)$ marked as "1" are used in the first-order noncausal linear prediction and shown in Figure. 1. The prediction coefficients $a_1(0, 1)$ and $a_1(1, 0)$ the first-order noncausal predictor are related with the autocorrelation functions by the following normal equation in Eq. 6: $R_{x_1} a_1 = r_{x_1}$ where

$$R_{x_1} = \begin{bmatrix} R_x(0, 0) + R_x(0, 2) & R_x(1, -1) + R_x(1, 1) \\ R_x(1, -1) + R_x(1, 1) & R_x(0, 0) + R_x(2, 0) \end{bmatrix} \text{ and } r_{x_1} = \begin{bmatrix} R_x(0, 1) \\ R_x(1, 0) \end{bmatrix}$$

The number of different elements from the autocorrelation matrix is 7 for the first-order noncausal predictor.

3.3 Second-order Causal Linear Predictor

The predicted pixel $\hat{x}(m, n)$ of the centre pixel $x(m, n)$ for the second-order causal linear prediction is given by

$$\begin{aligned}\hat{x}(m, n) &= a_2(0, 1)x(m, n-1) + a_2(0, 1)x(m-1, n) + a_2(1, 1)x(m-1, n-1) + a_2(1, -1)x(m+1, n-1) \\ &= x_2 a_2 = \begin{bmatrix} x(m, n-1) \\ x(m-1, n) \\ x(m-1, n-1) \\ x(m+1, n-1) \end{bmatrix} \begin{bmatrix} a_2(0, 1) \\ a_2(1, 0) \\ a_2(1, 1) \\ a_2(1, -1) \end{bmatrix} \quad (11)\end{aligned}$$

The pixels $x(m, n-1)$, $x(m-1, n)$, $x(m-1, n-1)$ and $x(m+1, n-1)$, marked as "1" and "2" are shown to the left side of dark line in Figure.1, and used in the second-order causal linear prediction. The prediction coefficients $a_2(0, 1)$, $a_2(1, 0)$, $a_2(1, 1)$ and $a_2(1, -1)$ for the second-order causal predictor are related with the autocorrelation functions by the following normal equation in Eq. 6: $R_{x_2} a_2 = r_{x_2}$ where

$$R_{x_2} = \begin{bmatrix} R_x(0, 0) & R_x(1, -1) & R_x(1, 0) & R_x(1, -2) \\ R_x(1, -1) & R_x(0, 0) & R_x(0, 1) & R_x(0, 1) \\ R_x(1, 0) & R_x(0, 1) & R_x(0, 0) & R_x(0, 2) \\ R_x(1, -2) & R_x(0, 1) & R_x(0, 2) & R_x(0, 0) \end{bmatrix} \text{ and } r_{x_2} = \begin{bmatrix} R_x(0, 1) \\ R_x(1, 0) \\ R_x(1, 1) \\ R_x(1, -1) \end{bmatrix}$$

The number of different elements from the autocorrelation matrix is 7 for the second-order causal predictor.

3.4 Second-order Noncausal Linear Predictor

The predicted value $\hat{x}(m, n)$ of the centre pixel $x(m, n)$ for the second-order noncausal linear prediction is given by [17]

$$\begin{aligned} \hat{x}(m, n) = & a_2(0, 1)x(m, n-1) + a_2(0, -1)x(m, n+1) + a_2(1, 0)x(m-1, n) + a_2(-1, 0)x(m+1, n) + \\ & a_2(1, 1)x(m-1, n-1) + a_2(-1, -1)x(m+1, n+1) + a_2(1, -1)x(m-1, n+1) + a_2(-1, 1) \\ & x(m+1, n-1) = a_2(0, 1)\{x(m, n-1) + x(m, n+1)\} + a_2(1, 0)\{x(m-1, n) + x(m+1, n)\} + \\ & a_2(1, 1)\{x(m-1, n-1) + x(m+1, n+1)\} + a_2(1, -1)\{x(m-1, n+1) + x(m+1, n-1)\} \\ & = x_2 a_2 \end{aligned} \quad (12)$$

where $a_2(0, 1) = a_2(0, -1)$, $a_2(1, 0) = a_2(-1, 0)$, $a_2(1, 1) = a_2(-1, -1)$ and $a_2(1, -1) = a_2(-1, 1)$ due to Eq. 7. The pixels $(m, n-1)$, $x(m, n+1)$, $x(m-1, n)$, $x(m+1, n)$, $x(m-1, n-1)$, $x(m+1, n+1)$, $x(m-1, n+1)$ and $x(m+1, n-1)$, marked as "1" and "2" are used in the second-order noncausal linear prediction and shown in Figure.1. The prediction coefficients $a_2(0, 1)$, $a_2(1, 0)$, $a_2(1, 1)$ and $a_2(1, -1)$ for the second-order noncausal predictor are related with the autocorrelation functions by the following normal equation in Eq. 6: $R_{x_2} a_2 = r_{x_2}$ where

$$R_{x_2} = \begin{bmatrix} R_x(0, 0) + R_X(0, 2) & R_x(1, -1) + R_X(1, 1) & R_x(1, 0) + R_X(1, 2) & R_x(1, -2) + R_X(1, 0) \\ R_x(1, -1) + R_X(1, 1) & R_x(0, 0) + R_X(2, 0) & R_x(0, 1) + R_X(2, 1) & R_x(0, 1) + R_X(2, -1) \\ R_x(1, 0) + R_X(1, 2) & R_x(0, 1) + R_X(2, 1) & R_x(0, 0) + R_X(2, 2) & R_x(0, 2) + R_X(2, 0) \\ R_x(1, -2) + R_X(1, 0) & R_x(0, 1) + R_X(2, -1) & R_x(0, 2) + R_X(2, 0) & R_x(0, 0) + R_X(2, -2) \end{bmatrix}$$

$$\text{and } r_{x_2} = [R_x(0, 1) \quad R_x(1, 0) \quad R_x(1, 1) \quad R_x(1, -1)]'$$

The number of different elements from the autocorrelation matrix is 13 for the second-order noncausal predictor. A block of the size $B \times B$ gives an autocorrelation matrix of the size $(2B-1) \times (2B-1)$. Selective elements for the autocorrelation matrix are given below. $R_X(0, 0) = R_{X_2}(B, B)$, $R_X(0, 2) = R_{X_2}(B, B-2)$, $R_X(1, -1) = R_{X_2}(B-1, B+1)$, $R_X(1, 1) = R_{X_2}(B-1, B-1)$, $R_X(1, 0) = R_{X_2}(B-1, B)$, $R_X(2, 0) = R_{X_2}(B-2, B)$, $R_X(0, 1) = R_{X_2}(B, B-1)$, $R_X(1, -2) = R_{X_2}(B-1, B+2)$, $R_X(1, 2) = R_{X_2}(B-1, B-2)$, $R_X(2, 1) = R_{X_2}(B-2, B-1)$, $R_X(2, -1) = R_{X_2}(B-2, B+1)$, $R_X(2, 2) = R_{X_2}(B-2, B-2)$, $R_X(2, -2) = R_{X_2}(B-2, B+2)$. The first 7 elements in the above are used to predict all pixels in that blocks for the first-order noncausal prediction, and all elements in the above to predict all pixels in that block of the image for the second-order noncausal linear prediction.

4 Experimental Results

A set of 100 high-quality RGB images was collected from the Internet including satellite images, biomedical images, texture images, aerial images, photographic images, computer generated images and classic images commonly used in the color image processing literature. All images are of size 512×512 Figure. 2 shows representative images from this set. The performance of the proposed algorithm depends on the following various parameters and results are taken based on these parameters.

- (I) Block-size of the prediction window
- (II) Noncausal or causal
- (III) order of prediction
- (IV) Value of k

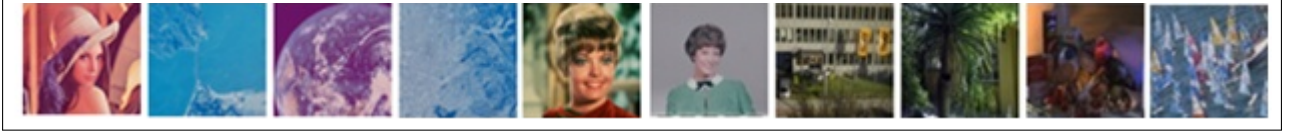


Figure 2: (a) Lena, (b) Golden Gate, (c) Earth, (d) Oakland, (e) Zelda, (f) Girl, (g) Building, (h) Aptus, (i) Hdr and (j) Sail.

4.1 Comparison of Performance based on Block-size

Experiments were taken to find the most appropriate block-size of the prediction using 100 different images. Results were taken for both the first-order causal and noncausal for the block-sizes of 4×4 , 8×8 , 16×16 , 32×32 and 64×64 respectively. These results in terms of percentage estimated length (EL) and percentage error (ER) are shown in Table 1 and Table 2. Results were taken based on the averages of EL and ER of 100 images for $k=9$. The results in these tables are shown for the first-order. It was found that the performance is poor for both smaller and larger block-sizes. The most appropriate block-size is 16×16 , which gives the least ER for majority of embedding percentages for both causal and noncausal predictors. The poor performance of the smaller block-size is due to the fact that the prediction coefficients obtained from smaller blocks can give accurate prediction if the blocks contain highly similar pixel values, but a block does not contain similar pixel values usually. On the other hand, larger block-sizes also give erroneous results, because the larger blocks contain many heterogeneous pixel values. Without loss of generality, block-size of 16×16 was used for all results afterward.

Table 1: Selection of the best block-size in terms of EL and ER for the first-order noncausal predictor

Block-size	Embedding Percentage										
	EL						ER				
	0%	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
4×4	-0.17	9.78	21.58	28.99	40.40	47.93	2.20	-15.80	10.10	-4.00	20.70
8×8	-0.64	9.77	22.83	28.88	39.56	50.58	2.30	-28.30	11.20	4.40	-5.80
16×16	-1.46	10.64	20.75	30.02	40.28	49.62	-6.40	-7.50	-0.20	-2.80	3.80
32×32	-0.94	10.13	23.50	29.69	40.28	46.60	-1.30	-35.00	3.10	-2.80	34.00
64×64	2.87	19.18	16.41	39.10	36.35	41.81	-91.80	35.90	-91.00	36.50	81.90

4.2 Comparison of Performance of Noncausal and Causal Predictors

Table 1 and Table 2 show that the noncausal predictors give less ER than causal predictor. The noncausal predictor uses all past and future pixels surrounding the pixels to be predicted and the causal predictor uses only the past pixels surrounding the pixel to be predicted. This results in better performance for the noncausal predictor.

4.3 Comparison of Performance based on Orders of Prediction

Table 3 gives the comparison of performance based on the orders of prediction. Results were based on the average results of all images for the causal and noncausal predictors with $k=9$. ER

Table 2: Selection of the best block-size in terms of EL and ER for the first-order causal predictor

Block-size	Embedding Percentage										
	EL						ER				
	0%	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
4×4	-2.51	12.53	23.03	28.63	38.95	49.22	-25.30	-30.30	13.70	10.50	7.80
8×8	-1.17	10.86	22.04	28.70	39.80	49.48	-8.60	-20.40	13.00	2.00	5.20
16×16	-1.58	10.66	21.16	30.49	39.92	48.77	-6.60	-11.60	-4.90	0.80	12.30
32×32	-1.82	10.84	21.85	30.61	39.94	48.49	-8.40	-18.50	-6.10	0.60	15.10
64×64	-1.80	11.10	21.33	30.75	39.87	48.56	-11.00	-13.30	-7.50	1.30	14.40

of the first-order predictor is comparatively less than that of the second-order predictor. Results of orders higher than the second are not shown, because performance deteriorates as the order increases.

Table 3: Performance based on different images for the first-order noncausal and causal second-order predictors in term of EL and ER

Type	Embedding Percentage											
	First-order noncausal linear predictor						Second-order noncausal linear predictor					
	0%	10%	20%	30%	40%	50%	0%	10%	20%	30%	40%	50%
EL	-1.46	10.64	20.75	30.02	40.28	49.62	-3.23	12.08	23.11	31.34	39.55	47.07
ER	—	-6.40	-7.50	-0.20	-2.80	3.80	—	-20.80	-15.55	-4.46	1.12	5.86

4.4 Performance for Different k

Table 4 gives the results of dependency of the performance of extraction on the values of k for noncausal and causal predictors on average results. The result is taken for values of k from 1 to 9 with the increment 2 for block size of 16×16 . Tables show that the performance of the proposed technique was better for higher values of k . It is because the predictor can give larger value of absolute error $|e(m, n)|$ when the predicted pixel is different from the neighboring pixels.

Table 4: Performance based on value of k for the first-order noncausal predictor in terms of EL and ER

Value of k	Embedding Percentage										
	EL						ER				
	0%	10%	20%	30%	40%	50%	10%	20%	30%	40%	50%
1	18.70	23.73	17.08	29.11	25.34	28.21	-137.30	29.20	8.90	146.60	217.90
3	1.92	11.27	23.32	32.45	31.11	52.08	-12.70	-33.20	-24.50	88.90	-20.80
5	1.48	12.83	21.57	36.55	37.23	48.81	-28.30	-15.70	-65.50	27.70	11.90
7	1.78	9.10	18.62	31.88	38.16	49.45	9.00	13.80	-18.80	18.40	5.50
9	-1.46	10.64	20.75	30.02	40.28	49.62	-6.40	-7.50	-0.20	-2.80	3.80

5 Conclusions

This paper presents a novel steganalytic technique of $\pm k$ steganography based on noncausal linear predictor using prediction coefficients obtained from a block of pixels in the image. The effectiveness of the proposed technique in estimating the length of the embedded message bits in the stego-image was verified using different types of images. It was found that noncausal predictor gives better results than causal predictor, and the first-order predictor gives better performance than the second-order. It is also found that the performance is better for higher values of k .

Bibliography

- [1] Kawaguchi E.; Noda H.; Niimi M.; Eason R.O. (2003); A Model of Anonymous Covert Mailing System Using Steganographic Scheme, in *Information modelling and knowledge bases*, XIV, H. Yaakkola et al. (Eds), IOS Press, 81-85.
- [2] Wang H.; Wang S. (2004); Cyber Warfare Steganography vs Steganalysis, *ACM Commun.*, 47: 76-82.
- [3] Anderson R.J.; Pettitcolas, F.A.P. (1998); On The Limits Of Steganography, *IEEE Journal on Selected Areas in Communication*, 16(4): 474-481.
- [4] Zhang X.; Wang S. (2006); Dynamically Running Coding In Digital Watermarking, *IEEE Signal Processing Letters*, 13(3): 165-168.
- [5] Provos N.; Honeyman P. (2003); Hide And Seek: An Introduction To Steganography, *IEEE Security And Privacy*, 1(3): 32-44.
- [6] Westfeld A.; Pfitzmann A. (1999); Attack On Steganographic Systems, *Proc. Information Hiding - 3rd Intel Workshop*, Springer Verlag, 61-76.
- [7] Fridrich J.; Goljan M. (2000); Practical Steganalysis Of Digital Images - State Of The Art, Security And Watermarking Of Multimedia Contents IV, E.J. Delp III and P.W. Wong, editors, *Proc. of SPIE*, 4675, 1-13.
- [8] Zhi L.; Fen S.A.; Xian Y.Y.A. (2003); LSB Steganography Detection Algorithm, *Proc. IEEE ISPIE*, 2780-2783.
- [9] Zhang T.; Ping X. (2003) ; Reliable Detection Of LSB Steganography Based On The Difference Image Histogram, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 3: 545-548.
- [10] Mielikainen J. (2006); LSB Matching Revisted, *IEEE Signal Processing Letters*, 13(5): 285-287.
- [11] Li X.; Yang B.; Cheng D.; Zeng T. (2009); A Generalization Of LSB Matching, *IEEE Signal Processing Letters*, 16(2): 69-72.
- [12] Fridrich J.; Soukal D.; Goljan M. (2005); Maximum Likelihood Destination Of Secret Message Length Embedded Using PMK Steganography In Spatial Domain, *Proc. Of IST/SPIE Electronic Imaging: Security, Steganography And Watermarking Of Multimedia Contents*, VII, 5681, 595-606.

- [13] Balram N.; Moura M.F. (1996); Noncausal Predictive Image Codec, *IEEE Trans. On Image Processing*, 5(8): 1229-1242.
- [14] Schroeder M.R.; Atal B.S. (1985); Code-excited Linear Prediction (CELP): High Quality Speech At Very Low Bit Rates, *IEEE Proc. ICASSP*, 937-940.
- [15] Fang W.; Yagle A.E. (1994); Two-dimensional Linear Prediction And Spectral Estimation In Polar Raster, *IEEE Trans. On Signal Processing*, 42(3): 628-641.
- [16] Jain A.K. (1989); *Fundamental of Digital Image Processing*, Prentice Hall, Information and System Science Series.
- [17] Singh K. M. (2012); Vector Median Filter Based On Noncausal Linear Prediction For Detection Of Impulse Noise From Images, Inderscience *International Journal of Computational Science and Engineering*, 7(4): 345-355.